

On the Achievable Rate of Bandlimited Continuous-Time AWGN Channels with 1-Bit Output Quantization

Sandra Bender, Meik Dörpinghaus, and Gerhard Fettweis

Abstract

We consider a continuous-time bandlimited additive white Gaussian noise channel with 1-bit output quantization. On such a channel the information is carried by the temporal distances of the zero-crossings of the transmit signal. The set of input signals is constrained by the bandwidth of the channel and an average power constraint. We derive a lower bound on the capacity by lower-bounding the achievable rate for a given set of waveforms with exponentially distributed zero-crossing distances. We focus on the behaviour in the high signal-to-noise ratio regime and characterize the achievable rate over the available bandwidth and the signal-to-noise ratio.

I. INTRODUCTION

For the design of digital communication systems, we typically assume that the analog-to-digital converter (ADC) at the receiver provides a sufficiently fine grained quantization of the magnitude of the received signal such that its quantization resolution is not the limiting factor regarding the achievable data rate. However, for very high data rate short link communication the power consumption of the ADC becomes a major factor, also in comparison to the transmit power. The combination of high quantization resolution with a very high sampling rate which is required when considering transmission bandwidths of up to 25 GHz leads to a high ADC power consumption. One idea to circumvent this problem is the usage of 1-bit channel output quantization in combination with oversampling the received signal with respect to the Nyquist rate. One-bit quantization is fairly simple to realize as the quantizer merely becomes a simple comparator without the need for highly linear analog signal processing. Using such an approach quantization resolution of the signal magnitude is traded-off by quantization resolution of the received signal in time domain. Obviously, optimal communication over the resulting channel requires a modulation and signaling scheme adapted to this specific channel as the information is no longer carried in the signal magnitude but in the zero-crossing time instants of the transmitted signal. The question then is, how much the channel capacity is degraded by such 1-bit quantized oversampled channel compared to an additive white Gaussian noise (AWGN) channel quantized with high resolution and sampled at Nyquist rate. In [1] simulative approaches on bounding the achievable rate in a discrete-time scenario are studied. In [2]–[4], the achievable rate is evaluated via simulation for different signaling strategies.

However, an analytical evaluation of the channel capacity of the 1-Bit quantized oversampled AWGN channel is still open but important to understand its fundamental behaviour. This capacity depends on the oversampling factor, as due to the 1-bit quantization Nyquist-sampling, like any other sampling, does not provide a sufficient statistic. As a limiting case, in the present work, we study the capacity of the underlying continuous-time 1-bit quantized channel, i.e., without any time discretization. The capacity of this channel corresponds to the limit of the capacity of the additive noise channel with 1-bit quantization and oversampling in case the sampling rate becomes infinitely large. Without time quantization, there is, as for the capacity of the AWGN channel as given by Shannon [5], no quantization in the information carrying dimension. The capacity of the AWGN without output quantization is obviously an upper bound on the capacity of the continuous-time 1-bit quantized channel. This approach also enables a better understanding of the difference between using the magnitude domain versus the time domain for signaling. As the continuous-time additive noise channel with 1-bit output quantization carries the information in the zero-crossings of the transmit signal, this channel corresponds to some extent to a timing channel as, e.g., studied in [6].

In general, the idea of increasing the information rate of a continuous-time channel with binary output quantization by oversampling is not new. Already in the early works by Gilbert [7] and Shamai [8] it has been shown that oversampling can increase the information rate, where in the latter paper the capacity has been lower-bounded by $\log_2(n + 1)$ [bits/Nyquist interval] where n is the oversampling rate with respect to sampling at Nyquist rate. However, both, [7] as well as [8], consider the noise free case exclusively. Differently, in [9] Koch and Lapidoth have shown that oversampling increases the capacity per unit-cost of bandlimited Gaussian channels with 1-bit output quantization. This work focuses on the low signal-to-noise ratio (SNR) domain. In [9] it has been shown that oversampling increases the achievable rate, here based on the study of

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center SFB 912 Highly Adaptive Energy-Efficient Computing (HAEC) and by the DFG in the framework of the cluster of excellence EXC 1056 Center for Advancing Electronics Dresden (cfaed).

S. Bender, M. Dörpinghaus, and G. Fettweis are with the Vodafone Chair Mobile Communications Systems and with the Collaborative Research Center SFB 912 HAEC, Technische Universität Dresden, 01062 Dresden, Germany (e-mail: sandra.bender, meik.doerpinghaus, fettweis@tu-dresden.de).

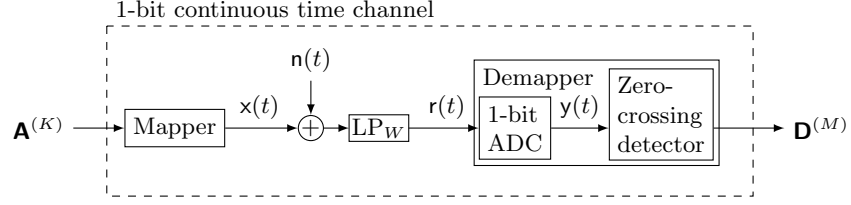


Fig. 1. System model

the generalized mutual information. All of the previously listed works considering noisy channels at the same time consider sampling with some fixed oversampling rate and do not consider the capacity of the underlying continuous-time channel.

Differently, in the present work we are interested in the capacity of the continuous-time channel. Given the outlined application scenario of sort range multigigabit/s-communication, we focus on the high SNR domain. We derive a lower bound on the capacity of the bandlimited continuous-time additive Gaussian noise channel with 1-bit output quantization. We show that the achievable rate increases with the bandwidth for an appropriately chosen input distribution but saturates over the SNR. Moreover, we observe that the ratio between our lower bound and the AWGN capacity is a constant independent of the bandwidth for a given SNR and the appropriately chosen input distribution mentioned before.

The rest of the paper is organized as follows. In Section II, the system model is introduced. In Section IV the different types of possible error events are discussed, which are modeled and analyzed separately in Sections IV and V. Subsequently, in Section VI the effects of the distortion introduced by the receive filter are discussed and Section VII concludes the paper with a summary of the results and a brief discussion.

II. SYSTEM MODEL AND DESIGN PARAMETERS

We consider the system model depicted in Fig. 1. As mentioned, a receiver relying on 1-bit quantization can only distinguish between the level of the input signal being smaller or larger than zero. Hence, all information that can be conveyed through such a channel is encoded in the time instants of the zero-crossings¹. In order to model this, we consider the channel input and output vectors, $\mathbf{A}^{(K)} = [A_1, \dots, A_K]^T$ and $\mathbf{D}^{(M)} = [D_1, \dots, D_M]^T$, which contain the temporal distances by A_k and D_m of two consecutive zero-crossings (ZC) of the transmit signal $x(t)$ and the received signal $r(t)$, respectively. Here K is not necessarily equal to M as noise can add or remove zero-crossings. For an infinite observation interval $K \rightarrow \infty$ and $M \rightarrow \infty$, we omit the superscript and denote the corresponding random processes \mathbf{A} and \mathbf{D} . For the analysis in this work, it is assumed that the time instants of the zero crossings can be resolved with infinite precision, which makes A_k and D_m continuous random variables. The mapper converts the random vector $\mathbf{A}^{(K)}$ into the continuous-time transmit signal $x(t)$ with limited degrees of freedom in amplitude, which is transmitted over a additive white Gaussian noise (AWGN) channel. At the receiver, lowpass-filtering with one-sided bandwidth W ensures bandlimitation as well in the signal as the noise. After lowpass-filtering the demapper realizes the conversion between the noisy received signal $r(t)$ and the sequence $\mathbf{D}^{(M)}$ of zero-crossing distances.

A. Signal Structure and Input Distribution

The input symbols A_k correspond to the temporal distances between the k th and the $(k-1)$ th zero-crossing in the transmit signal $x(t)$. We consider i.i.d. exponentially distributed A_k with

$$A_k \sim \lambda e^{-\lambda(a-\beta)} \mathbb{1}_{[\beta, \infty)}(a) \quad (1)$$

since the exponential distribution maximizes the entropy for positive continuous random variables with given mean. Here, $\mathbb{1}_{[u, v]}(x)$ is the indicator function, defined to be one in the interval $[u, v]$ and zero outside. This results in a mean symbol duration of

$$T_{\text{avg}} = \frac{1}{\lambda} + \beta. \quad (2)$$

Fig. 2 illustrates the mapping of the input sequence $\mathbf{A}^{(K)}$ to the transmit signal $x(t)$, which alternates between two levels $\pm\sqrt{\hat{P}}$, where \hat{P} is the peak power of the input signal. With β being the minimal value of the A_k , it is guaranteed that $x(t)$ achieves $\sqrt{\hat{P}}$ between two transitions. This is not necessarily capacity-achieving but simplifies the derivation of a lower bound on the capacity. The k th zero-crossing corresponding to A_k occurs at time

$$T_k = \sum_{i=1}^k A_i + t_0. \quad (3)$$

¹Note that one additional bit is carried by the sign of the signal. However, its effect on the mutual information between channel input and output can be neglected when studying the capacity as it converges to zero for infinite blocklength.

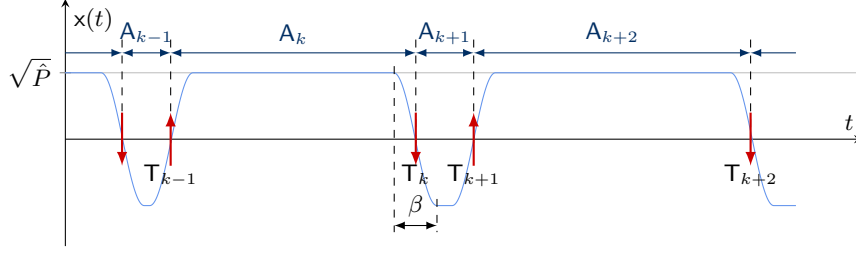


Fig. 2. Mapping from input sequence $\mathbf{A}^{(K)}$ to transmit signal $x(t)$

Without loss of generality, we assume $t_0 = 0$. In order to control the bandwidth of the input signal and match it to the channel, the transition from one level to the other is not a step function but modeled by the waveform $f(t)$, yielding the transmit signal

$$x(t) = \left(\sum_{k=1}^K \sqrt{\hat{P}} (-1)^k g(t - T_k) \right) + \sqrt{\hat{P}} \quad (4)$$

with the pulse shape

$$g(t) = \left(1 + f\left(t - \frac{\beta}{2}\right) \right) \cdot \mathbb{1}_{[0, \beta]} + 2 \cdot \mathbb{1}_{[\beta, \infty)} \quad (5)$$

where $f(t)$ is an odd function between $(-\beta/2, -1)$ and $(\beta/2, 1)$, describing the transition of the signal. The transition time β is chosen according to the available bandwidth W of the channel with

$$\beta = \frac{1}{2W}. \quad (6)$$

If not stated otherwise, results throughout the paper are given for a sine halfwave as transition, i.e.,

$$f(t) = \begin{cases} \sin\left(\pi \frac{t}{\beta}\right) & \text{for } |t| \leq \beta/2 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

yielding

$$g(t) = \left(1 - \cos\left(\pi \frac{t}{\beta}\right) \right) \cdot \mathbb{1}_{[0, \beta]} + 2 \cdot \mathbb{1}_{[\beta, \infty)}. \quad (8)$$

In the limiting case of $\lambda \rightarrow \infty$, i.e., very short input symbols and very frequent transitions, c.f. (1), this leads to a one sided signal bandwidth of W . Hence, we can consider the signal $x(t)$ as almost bandlimited, where only a small portion of the signal energy is outside the interval $[-W, W]$. Strict bandlimitation of the channel is ensured by the lowpass-filter before the demapping.

B. Channel Model

We consider a continuous-time additive Gaussian noise channel with 1-bit output quantization. The received signal after lowpass-filtering is given by

$$r(t) = \hat{x}(t) + \hat{n}(t) \quad (9)$$

where $\hat{x}(t)$ and $\hat{n}(t)$ denote the filtered version of the transmit signal and the noise, respectively. The receive filter is considered to be an ideal lowpass with one-sided bandwidth W . After quantization the signal fed to the decoder is

$$y(t) = Q(r(t)) \quad (10)$$

where Q denotes a binary quantizer with a threshold at zero and is thus defined as

$$Q(s) = \begin{cases} 1 & \text{if } s \geq 0 \\ -1 & \text{otherwise} \end{cases} \quad (11)$$

As the transmitted signal is not strictly bandlimited, for finite channel bandwidths W it holds that $x(t) \neq \hat{x}(t)$. We model this by introducing a filter distortion

$$\tilde{x}(t) = \hat{x}(t) - x(t), \quad (12)$$

which we treat as additional noise source. Thus, we can write for the received signal

$$r(t) = x(t) + z(t) \quad (13)$$

where

$$z(t) = \hat{n}(t) + \tilde{x}(t) \quad (14)$$

is the overall distortion introduced by the noise and the lowpass-filtering. The noise $n(t)$ is zero-mean additive white Gaussian noise with power spectral density (PSD) $N_0/2$. Its filtered version is $\hat{n}(t)$ has the PSD

$$S_{\hat{n}}(f) = \begin{cases} \frac{N_0}{2} & \text{for } |f| \leq W \\ 0 & \text{otherwise} \end{cases}. \quad (15)$$

Accordingly, the variance of the Gaussian noise is

$$\sigma_{\hat{n}}^2 = \int_{-W}^W \frac{N_0}{2} df = N_0 W \quad (16)$$

and the variance of $\tilde{x}(t)$ is given by

$$\sigma_{\tilde{x}}^2 = \mathbb{E} \left[|x(t) - \hat{x}(t)|^2 \right]. \quad (17)$$

C. System Parameters and their Dependencies

In the following, the key system parameters are defined and their relations are described. The signal-to-noise ratio is defined as

$$\rho = \frac{P}{N_0 W}, \quad (18)$$

where P is the average transmit signal power, i.e., $P = \lim_{K \rightarrow \infty} \frac{1}{K T_{\text{avg}}} \mathbb{E} \left[\int_0^T |x(t)|^2 dt \right]$. It depends on the transition time β , the average symbol duration T_{avg} and the waveform $f(t)$ and is given by

$$P = \frac{\hat{P}}{T_{\text{avg}}} \left(\int_0^\beta f^2 \left(t - \frac{\beta}{2} \right) dt + \lambda^{-1} \right). \quad (19)$$

For the cosine waveform in (7) it results

$$\begin{aligned} P &= \frac{\hat{P}}{T_{\text{avg}}} \left(\int_0^\beta \cos^2 \left(\frac{\pi}{\beta} t \right) dt + \lambda^{-1} \right) \\ &= \frac{\frac{1}{2} + 2W\lambda^{-1}}{1 + 2W\lambda^{-1}} \hat{P} \end{aligned} \quad (20)$$

which shows that the SNR depends on the peak power \hat{P} , the bandwidth W , the noise power N_0 , the input distribution in form of λ , and the waveform $f(t)$. Those five parameters can be varied independently of each other. The SNR, the signal spectrum, the distortion variance σ_z^2 etc. depend on those parameters.

III. ACHIEVABLE RATE AND ERROR EVENTS

The capacity of a communication channel represents the highest rate that can be used to transmit over the channel with an arbitrary small probability of error and is defined as

$$C = \sup I'(\mathbf{A}; \mathbf{D}) \quad (21)$$

where the supremum is taken over all distributions of the input signal for which the set of all input signals has a constrained average power P and a constrained bandwidth W . In (21) the achievable rate is given by

$$I'(\mathbf{A}; \mathbf{D}) = \lim_{K \rightarrow \infty} \frac{1}{K T_{\text{avg}}} I(\mathbf{A}^{(K)}; \mathbf{D}^{(M)}) \quad (22)$$

with $I(\mathbf{A}^{(K)}; \mathbf{D}^{(M)})$ being the mutual information. Note that we have defined the mutual information rate based on a normalization with respect to the expected transmission time $K T_{\text{avg}}$. In the present paper, we do not consider the evaluation of the supremum over the input distribution, but restrict ourselves to input signal as described in Section II-A, yielding a lower bound on the capacity. However, later we will consider the supremum of $I'(\mathbf{A}; \mathbf{D})$ over the parameter λ of the distribution of the A_k in (1). In order to evaluate $I'(\mathbf{A}; \mathbf{D})$, we have to study, how the noise and the lowpass-filtering alter the zero-crossings of the received signal $r(t)$ w.r.t. $x(t)$. Zero-crossings can be shifted, leading to an error in magnitude of A_k , or a pair of zero-crossings can be either introduced or deleted, leading to insertion or deletion of symbols. Hence, the channel is an additive noise channel with insertions and deletions. For such channels are to the best of our knowledge only results for binary channels in form of combinatorial solutions available, e.g., [10]–[13]. For the considered input signals and high SNR

scenario, the occurrence of deletions is assumed to be negligible since $A_k \geq \beta$ and β depends directly on the bandwidth of the receiver. Hence, the filtered noise can be assumed to be uncorrelated in an interval longer as β and the possibility of a noise event inverting an entire symbol can be neglected.

The remaining error events, shift and insertions of zero-crossing, can be analyzed separately using the idea of a genie-aided receiver as in [12]. We introduce an auxiliary process \mathbf{V} providing information about the inserted zero-crossings to the receiver, such that it can remove the additional zero-crossings. This process will be described below. Let $\hat{\mathbf{D}}$ contain the temporal distances of the zero-crossings at the receiver when the additional zero-crossings are removed. Hence, $\hat{\mathbf{D}}$ can be calculated based on \mathbf{D} and \mathbf{V} and it holds for the mutual information rate in case the receiver has the side information about the inserted zero-crossings

$$I'(\mathbf{A}; \hat{\mathbf{D}}) = I'(\mathbf{A}; \mathbf{D}, \mathbf{V}). \quad (23)$$

Using the chain rule, we have

$$I'(\mathbf{A}; \mathbf{D}, \mathbf{V}) = I'(\mathbf{A}; \mathbf{D}) + I'(\mathbf{A}; \mathbf{V}|\mathbf{D}). \quad (24)$$

Thus

$$I'(\mathbf{A}; \mathbf{D}) = I'(\mathbf{A}; \mathbf{D}, \mathbf{V}) - I'(\mathbf{A}; \mathbf{V}|\mathbf{D}) \quad (25)$$

where $I'(\mathbf{A}; \mathbf{D})$ is the mutual information rate without side information at the receiver. Hence, $I'(\mathbf{A}; \mathbf{V}|\mathbf{D})$ describes the reduction of the mutual information rate due to not knowing \mathbf{V} . This approach allows to evaluate separately the effect of the shifted zero-crossings, captured in $I'(\mathbf{A}; \mathbf{D}, \mathbf{V})$, and the influence of the inserted zero-crossings, described by $I'(\mathbf{A}; \mathbf{V}|\mathbf{D})$.

For the characterization of the auxiliary process \mathbf{V} , we consider for the moment the transmission of one input symbol A_k . Its bounding zero-crossings T_{k-1} and T_k will be shifted to \hat{T}_{k-1} and \hat{T}_k by the noise process, such that

$$\hat{T}_k = T_k + S_k \quad (26)$$

where S_k is a shift in time caused by the overall noise $z(t)$. The corresponding process will be denoted \mathbf{S} . Furthermore, additionally introduced zero-crossings will divide the input symbol into a vector of corresponding received symbols. The latter is reversible, if the receiver knows which zero-crossings correspond to the originally transmitted ones. The receiver would need to sum up the distances D_m that are separated by the additional zero-crossings in order to obtain the corresponding symbols \hat{D}_k . Intuitively, one would start such an algorithm with the first received symbol, which gives way to the following thought: Instead of providing the receiver with the exact positions in time of the additional zero-crossings, it is sufficient to know - from the first transmitted symbol onwards - how many received symbols have to be summed up to generate sequence $\hat{\mathbf{D}}^{(K)}$. Hence, the auxiliary sequence $\mathbf{V}^{(K)}$ consists of positive integer numbers $V_k \in \mathbb{N}$, representing for each input symbol the number of corresponding output symbols. Thus, the auxiliary process \mathbf{V} is discrete, which we use for bounding the information rate in (25) by

$$\begin{aligned} I'(\mathbf{A}; \mathbf{D}) &= I'(\mathbf{A}; \mathbf{D}, \mathbf{V}) - H'(\mathbf{V}|\mathbf{D}) + H'(\mathbf{V}|\mathbf{D}, \mathbf{A}) \\ &\geq I'(\mathbf{A}; \mathbf{D}, \mathbf{V}) - H'(\mathbf{V}|\mathbf{D}) \end{aligned} \quad (27)$$

$$\geq I'(\mathbf{A}; \mathbf{D}, \mathbf{V}) - H'(\mathbf{V}) \quad (28)$$

where (27) results from the fact that the entropy rate of a discrete random process is non-negative and (28) is due to the fact that conditioning cannot increase entropy. The proof of the existence of a coding theorem remains for future research. In the following, we will derive bounds on $I'(\mathbf{A}; \mathbf{D}, \mathbf{V})$ and $H'(\mathbf{V})$.

IV. ACHIEVABLE RATE OF THE GENIE-AIDED RECEIVER

To evaluate the achievable rate $I'(\mathbf{A}; \hat{\mathbf{D}})$ of the genie-aided receiver, cf., (28) and (23), we have to evaluate the mutual information rate between the sequence of temporal spacings of zero-crossings at the channel input $\mathbf{A}^{(K)}$ and sequence of the temporal spacing of zero-crossings at the 1-bit quantizer $\hat{\mathbf{D}}^{(K)}$. Note, that in contrast to the original channel, here both vectors, $\mathbf{A}^{(K)}$ and $\hat{\mathbf{D}}^{(K)}$, are of same length as additional zero-crossings are removed at the receiver. The only error remaining is a shift S_k of every zero-crossings instant T_k to \hat{T}_k . Hence, on a symbols level we can write for the channel output

$$\begin{aligned} \hat{D}_k &= \hat{T}_k - \hat{T}_{k-1} \\ &= T_k + S_k - (T_{k-1} + S_{k-1}) \\ &= A_k + S_k - S_{k-1} \end{aligned} \quad (29)$$

The mutual information rate of this channel is then defined as

$$I'(\mathbf{A}; \hat{\mathbf{D}}) = \lim_{K \rightarrow \infty} \frac{1}{KT_{\text{avg}}} I(\mathbf{A}^{(K)}; \hat{\mathbf{D}}^{(K)}) \quad (30)$$

For the given system model, the following assumption are reasonable

(A1) the shifts S_k are independent

(A2) there is only one zero-crossing in each transition interval $\left[T_k - \frac{\beta}{2}, T_k + \frac{\beta}{2}\right]$

Assumption (A1) is easy to verify giving thought to the fact that any S_k and S_{k-1} are spaced at least time β apart, which is above the coherence time of the noise. Likewise, (A2) is based on the bandlimitation of the noise. Due to this, fast changes of the noise are impossible, such that additional zero-crossings within the transition interval are very unlikely. This has been verified by numerical computation based on curve-crossing problems for Gaussian random processes. The results are presented in Appendix A and show that this assumption is fulfilled for an SNR above 5 dB.

A. The Distribution of the Shifting Error

The distribution of S_k can be evaluated by mapping the probability density function (pdf) of the additive noise $z(T_k)$ at the time instant T_k by the function

$$z(T_k) = -\sqrt{\hat{P}}f(S_k) = -\sqrt{\hat{P}}\sin\left(\frac{\pi}{\beta}S_k\right) \quad (31)$$

into the zero-crossing error S_k on the time axis. The mapping hereby depends solely on the slope $f(t)$ of the transition waveform.

As $r(t)$ is almost bandlimited, it can be adequately described by a sampled representation with sampling rate $1/\beta$ to fulfill the Nyquist condition, cf. (6). Note that we here refer to the concept of sampling only to evaluate the value of $z(t)$ at the time instant T_k of the original zero-crossing. We still assume the receiver to be able to resolve the zero-crossings instants with infinite resolution.

For real valued random variables with a given variance, the Gaussian distribution is entropy maximizing. As we are deriving a lower bound on the achievable rate of the channel, we consider $z(t)$ to be Gaussian. The pdf of the additive Gaussian noise at the sampling time instant T_k , thus, is

$$p_z(z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left\{-\frac{z^2}{2\sigma_z^2}\right\} \quad (32)$$

with the variance σ_z^2 given by

$$\begin{aligned} \sigma_z^2 &= \sigma_{\hat{n}}^2 + \sigma_{\tilde{x}}^2 \\ &= N_0W + \sigma_{\tilde{x}}^2. \end{aligned} \quad (33)$$

as $\hat{n}(t)$ and $\tilde{x}(t)$ are independent. Hence, the probability density of S_k is given by

$$\begin{aligned} p_S(s) &= \left| \frac{\partial f(s)}{\partial s} p_z(f(s)) \right| \\ &= \sqrt{\frac{\pi\hat{P}}{2\sigma_z^2}} \frac{\cos\left(\frac{\pi}{\beta}s\right)}{\beta} \exp\left\{-\frac{\hat{P}}{2\sigma_z^2} \sin^2\left(\frac{\pi}{\beta}s\right)\right\}. \end{aligned} \quad (34)$$

As we are focusing on the high SNR behaviour of the capacity, the pdf $p_S(s)$ becomes narrow as the variance is given by $\frac{\sigma_z^2\beta^2}{\hat{P}\pi^2}$, such that only for $|s|$ relatively small $p_S(s)$ is significantly larger than zero. However, this means that for high SNR the zero-crossing errors S_k are with high probability small in comparison to the transition time β such that we can assume $s/\beta \ll 1$. In this case the pdf in (34) can be well approximated by

$$p_S(s) = \sqrt{\frac{\pi\hat{P}}{2\sigma_z^2}} \frac{1}{\beta} \exp\left\{-\frac{\hat{P}}{2\sigma_z^2} \left(\frac{\pi}{\beta}s\right)^2\right\}. \quad (35)$$

Hence, in the high SNR case, i.e., for SNRs above 6 dB, cf. Appendix B, the zero-crossing errors S_k can be approximated to be zero-mean Gaussian distributed with variance

$$\sigma_S^2 = \frac{\sigma_z^2}{4\pi^2W^2\hat{P}} \quad (36)$$

i.e., $S_k \sim \mathcal{N}(0, \sigma_S^2)$.

B. Lower Bound on the Achievable Rate of the Genie-Aided Receiver

For transmission blocks of a length corresponding to K zero-crossing intervals the input-output relation in (29) can be described by the following matrix-vector representation

$$\hat{\mathbf{D}}^{(K)} = \mathbf{A}^{(K)} + \mathbf{U}^{(K)} \mathbf{S}^{(K)} - \mathbf{S}_0^{(K)} \quad (37)$$

with

$$\mathbf{S}^{(K)} = [S_1, \dots, S_K]^T \quad (38)$$

$$\mathbf{S}_0^{(K)} = [S_0, 0, \dots, 0]^T \quad (39)$$

being both of length K . Here, S_0 denotes the zero-crossing error corresponding T_0 . Moreover, $\mathbf{U}^{(K)}$ is a $K \times K$ matrix of the following form

$$\mathbf{U}^{(K)} = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ -1 & 1 & 0 & \dots & \vdots \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix}. \quad (40)$$

The mutual information between the temporal spacings of the zero-crossings of the channel input signal $\mathbf{A}^{(K)}$ on the one hand, and the zero-crossings of the signal at the input to the 1-bit quantizer $\hat{\mathbf{D}}^{(K)}$ on the other hand, cf. (30), is given by

$$\begin{aligned} I(\mathbf{A}^{(K)}; \hat{\mathbf{D}}^{(K)}) &= h(\mathbf{A}^{(K)}) - h(\mathbf{A}^{(K)} | \hat{\mathbf{D}}^{(K)}) \\ &= h(\mathbf{A}^{(K)}) - h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} | \hat{\mathbf{D}}^{(K)}) \end{aligned} \quad (41)$$

where $h(\cdot)$ denotes the differential entropy. Moreover, $\hat{\mathbf{A}}_{\text{LMMSE}}^{(K)}$ is the linear minimum mean-squared error estimate of $\mathbf{A}^{(K)}$ based on $\hat{\mathbf{D}}^{(K)}$. Equality (41) follows from the fact that addition of a constant does not change differential entropy and the fact that $\hat{\mathbf{A}}_{\text{LMMSE}}^{(K)}$ can be treated as a constant while conditioning on $\hat{\mathbf{D}}^{(K)}$ as it is a deterministic function of $\hat{\mathbf{D}}^{(K)}$.

Next, we will upper-bound the second term on the RHS of (41), i.e., $h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} | \hat{\mathbf{D}}^{(K)})$. This term describes the randomness of the linear minimum mean-squared estimation error while estimating $\mathbf{A}^{(K)}$ based on the observation $\hat{\mathbf{D}}^{(K)}$. It can be upper-bounded by the differential entropy of a Gaussian random variable having the same covariance matrix [14, Theorem 8.6.5]. The estimation error covariance matrix of the linear minimum mean-squared error (LMMSE) estimator is given by

$$\begin{aligned} \mathbf{R}_{\text{err}}^{(K)} &= E \left[\left(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} \right) \left(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} \right)^T \right] \\ &= \sigma_A^2 \mathbf{I}^{(K)} - \sigma_A^4 \left(\sigma_A^2 \mathbf{I}^{(K)} + \sigma_S^2 \mathbf{R}_S^{(K)} \right)^{-1}. \end{aligned} \quad (42)$$

where

$$\begin{aligned} \sigma_A^2 \mathbf{I}^{(K)} &= E \left[\left(\mathbf{A}^{(K)} - \mu_A \right) \left(\mathbf{A}^{(K)} - \mu_A \right)^T \right] \\ &= \lambda^{-2} \mathbf{I}^{(K)} \end{aligned} \quad (43)$$

with

$$\begin{aligned} \mu_A &= E \left[\mathbf{A}^{(K)} \right] \\ &= (\beta + \lambda^{-1}) \mathbf{1}^{(K)} \end{aligned} \quad (44)$$

cf. (2) for (44), and where (43) follows from the fact that the elements of $\mathbf{A}^{(K)}$ are exponentially distributed, see (1). Furthermore, $\mathbf{I}^{(K)}$ is the identity matrix of size $K \times K$ and $\mathbf{1}^{(K)}$ is the all one column vector of length K . Moreover, $\sigma_S^2 \mathbf{R}_S^{(K)}$ is the covariance matrix of $\mathbf{U}^{(K)} \mathbf{S}^{(K)} - \mathbf{S}_0^{(K)}$, i.e.,

$$\begin{aligned} \sigma_S^2 \mathbf{R}_S^{(K)} &= \mathbb{E} \left[\left(\mathbf{U}^{(K)} \mathbf{S}^{(K)} - \mathbf{S}_0^{(K)} \right) \left(\mathbf{U}^{(K)} \mathbf{S}^{(K)} - \mathbf{S}_0^{(K)} \right)^T \right] \\ &= \sigma_S^2 \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix} \end{aligned} \quad (45)$$

of size $K \times K$. Here, we have used that $\mathbf{U}^{(K)} \mathbf{S}^{(K)} - \mathbf{S}_0^{(K)}$ is zero-mean.

Thus, the differential entropy $h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} | \hat{\mathbf{D}}^{(K)})$ is upper-bounded by

$$h(\mathbf{A}^{(K)} - \hat{\mathbf{A}}_{\text{LMMSE}}^{(K)} | \mathbf{D}^{(K)}) \leq \frac{1}{2} \log \det \left(2\pi e \mathbf{R}_{\text{err}}^{(K)} \right) \quad (46)$$

yielding the following lower bound for the mutual information in (41)

$$\begin{aligned} I(\mathbf{A}^{(K)}; \hat{\mathbf{D}}^{(K)}) &\geq h(\mathbf{A}^{(K)}) - \frac{1}{2} \log \det \left(2\pi e \mathbf{R}_{\text{err}}^{(K)} \right) \\ &= Kh(\mathbf{A}_k) + \frac{1}{2} \log \det \left((2\pi e)^{-1} \left(\sigma_A^{-2} \mathbf{I}^{(K)} + \sigma_S^{-2} (\mathbf{R}_S^{(K)})^{-1} \right) \right) \end{aligned} \quad (47)$$

where the first term of (47) follows from the independency of the elements of $\mathbf{A}^{(K)}$ and for the second term we have used (42) and the matrix inversion lemma.

With (47) the mutual information rate in (30) is lower-bounded by

$$\begin{aligned} I'(\mathbf{A}; \hat{\mathbf{D}}) &\geq \lim_{K \rightarrow \infty} \frac{1}{KT_{\text{avg}}} \left\{ Kh(\mathbf{A}_k) + \frac{1}{2} \log \det \left((2\pi e)^{-1} \left(\sigma_A^{-2} \mathbf{I}^{(K)} + \sigma_S^{-2} (\mathbf{R}_S^{(K)})^{-1} \right) \right) \right\} \\ &= \frac{1}{T_{\text{avg}}} \left\{ h(\mathbf{A}_k) + \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \log \left((2\pi e)^{-1} \left(\sigma_A^{-2} + \sigma_S^{-2} (S_S(f))^{-1} \right) \right) df \right\} \end{aligned} \quad (48)$$

where for (48) we have used Szegő's theorem on the asymptotic eigenvalue distribution of Hermitian Toeplitz matrices [15, pp. 64-65], [16] with $S_S(f)$ being the power spectral density corresponding to the sequence of covariance matrices $\mathbf{R}_S^{(K)}$. It is given by

$$\begin{aligned} S_S(f) &= -e^{j2\pi f} + 2 - e^{-j2\pi f} \\ &= 2(1 - \cos(2\pi f)), \quad |f| < 0.5. \end{aligned} \quad (49)$$

Although $S_S(f)$ is equal to zero for $f = 0$ it can be shown that the integral in (48) exists, as

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \log \left(\frac{\sigma_A^{-2}}{2\pi e} \left(1 + \frac{\sigma_A^2}{2\sigma_S^2(1 - \cos(2\pi f))} \right) \right) df = \log \left(\frac{\sigma_A^{-2}}{2\pi e} \right) + \text{arcosh} \left(\frac{\sigma_A^2}{2\sigma_S^2} + 1 \right). \quad (50)$$

As \mathbf{A}_k is exponentially distributed, we get

$$h(\mathbf{A}_k) = 1 - \log(\lambda). \quad (51)$$

With (2), (6), (36), (43), (50), and (51), the lower bound in (48) can be written as

$$\begin{aligned} I'(\mathbf{A}; \hat{\mathbf{D}}) &\geq \frac{1}{2T_{\text{avg}}} \left\{ \log \left(\frac{e}{2\pi} \right) + \text{arcosh} \left(\frac{1}{2\sigma_S^2 \lambda^2} + 1 \right) \right\} \\ &= \frac{W}{1 + 2W\lambda^{-1}} \left\{ \log \left(\frac{e}{2\pi} \right) + \text{arcosh} \left(\frac{2\pi^2 W^2 \hat{P}}{\sigma_z^2 \lambda^2} + 1 \right) \right\}. \end{aligned} \quad (52)$$

V. CHARACTERIZATION OF THE PROCESS OF ADDITIONAL ZERO-CROSSINGS

In order to lower-bound the rate $I'(\mathbf{A}; \mathbf{D})$ without the side information provided by \mathbf{V} to the receiver in Section IV, it remains to find an explicit expression or an upper bound for $H'(\mathbf{V})$ in order to derive a lower bound on the achievable rate, cf., (28). For every input symbol A_k the random variable V_k , which describes the number of received symbols that correspond to the transmitted one, depends on the number N_k of introduced zero-crossings by

$$V_k = N_k + 1. \quad (53)$$

Hence, we need to determine the number of times the distortion process $z(t)$ equals the negative transmit signal value within one symbol as then the received signal will be zero. Based on assumption (A2) we do not need to consider the transition intervals as they just contain the shifted zero-crossing. It remains the time $T_{\text{sat}} = \mathbb{E}[A_k] - \beta = \lambda^{-1}$ in which the signal level $\pm\sqrt{\hat{P}}$ is maintained, leading to a level-crossing problem. Level-crossing problems, especially for Gaussian processes, have been studied over decades, e.g., by Kac [17], Rice [18], Cramer and Leadbetter [19]. In order to be able to derive a closed-form expression for the lower bound on $I'(\mathbf{A}; \mathbf{D})$, we will derive a bound on $H'(\mathbf{V})$ based on the first moment of the distribution of V_k . The expected number of crossings of a level u in a time interval T for a Gaussian random process is given by the Rice formula [18]

$$\mathbb{E}[N_u(T)] = \frac{T}{\pi} \sqrt{\frac{-s''_{zz}(0)}{s_{zz}(0)}} \exp\left(-\frac{u^2}{2\psi_0}\right), \quad (54)$$

Here, $s_{zz}(\tau)$ is the autocorrelation function (ACF) of the Gaussian process $z(t)$ and $s''_{zz}(\tau) = \partial/\partial\tau^2 s_{zz}(\tau)$. In order for $\mathbb{E}[N_u(T)]$ to be finite, $-s''_{zz}(0) < \infty$ has to hold. Here, it holds that $u^2 = \hat{P}$, $T = \lambda^{-1}$, and $s_{zz}(0) = \sigma_z^2$ such that

$$\mu = \mathbb{E}[V_k] = \mathbb{E}[N_k] + 1 = \frac{1}{\pi} \sqrt{\frac{-s''_{zz}(0)}{\sigma_z^2}} \exp\left(-\frac{\hat{P}}{2\sigma_z^2}\right) \lambda^{-1} + 1. \quad (55)$$

Analogously to (33), we get

$$\begin{aligned} s''_{zz}(0) &= s''_{\hat{n}\hat{n}}(0) + s''_{\hat{x}\hat{x}}(0) \\ &= -\frac{4}{3}N_0W^3 + s''_{\hat{x}\hat{x}}(0). \end{aligned} \quad (56)$$

Using (55) we upper-bound the entropy rate $H'(\mathbf{V})$. For a given mean μ , the entropy maximizing distribution for a positive, discrete random variable is the geometric distribution, cf. [20, Section 1.9.32]. Hence, we can upper-bound the entropy $H(V_k)$ by

$$H(V_k) \leq (1 - \mu) \log(\mu - 1) + \mu \log \mu. \quad (57)$$

The derivation of (57) is given in Appendix C. All time intervals with maximum signal level $\pm\sqrt{\hat{P}}$ are spaced by transition time β apart, and, hence, all V_k can be considered to be i.i.d. - analogously to the S_k . This gives for the entropy rate of the auxiliary process

$$H'(\mathbf{V}) = \frac{1}{T_{\text{avg}}} H(V_k). \quad (58)$$

Note that the bound on $H(V_k)$ is an increasing function in μ . Furthermore, the expected number of level-crossings of a random Gaussian process increases with its variance. Hence, to evaluate (55), an upper bound for σ_z^2 and, thus, for $\sigma_{\hat{x}}^2$ is required. An upper bound on $\sigma_{\hat{x}}^2$ results into a lower bound on $s''_{\hat{x}\hat{x}}(0)$, cf. Section VI-B, as the two parameters depend on the ACF of the noise process and cannot be chosen independently. Both bounds will be derived in the next section.

VI. SIGNAL DISTORTION BY LOWPASS-FILTERING

The distortion of $x(t)$ introduced by the lowpass-filter can be quantified by the clipped energy, using the mean squared error $\sigma_{\hat{x}}^2$ as distortion measure, which is given by

$$\begin{aligned} \sigma_{\hat{x}}^2 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^T \mathbb{E}[\tilde{x}^2(t)] dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\tilde{x}}(\omega) d\omega, \end{aligned} \quad (59)$$

where (59) is Parseval's Theorem with $S_{\tilde{x}}(\omega)$ being the PSD of $\tilde{x}(t)$. As we consider a rectangular filter with cutoff-frequency W , it holds

$$S_{\tilde{x}}(f) = \begin{cases} S_x(f) & |f| > W \\ 0 & |f| \leq W \end{cases}, \quad (60)$$

with $S_X(f)$ being the PSD of the transmit signal $x(t)$. As $S_X(\omega)$ is even, we get

$$\sigma_x^2 = \frac{1}{\pi} \int_W^\infty S_X(\omega) d\omega \quad (61)$$

In order to evaluate (61), we derive $S_X(\omega)$. Steps on bounding the MSE, the correlation function and its derivation are taken subsequently.

A. Signal Spectrum and Autocorrelation Function

The power spectral density of a random process is defined as

$$S_X(\omega) = \lim_{K \rightarrow \infty} \frac{\mathbb{E} [|X(\omega)|^2]}{KT_{\text{avg}}} \quad (62)$$

where $X(\omega)$ is the spectrum of the random process $x(t)$ defined in (4) and given by

$$X(\omega) = \left(\sum_{k=1}^K \sqrt{\hat{P}}(-1)^k G(\omega) e^{-j\omega T_k} \right) + \sqrt{\hat{P}} 2\pi \delta(\omega) \quad (63)$$

where $G(\omega)$ is the Fourier transformation of the waveform $g(t)$ which is defined in (5). It holds that

$$G(\omega) = -j \left[\frac{1 + e^{-j\omega\beta}}{\omega} - e^{-j\omega\frac{\beta}{2}} a(\omega) \right], \quad (64)$$

where $a(\omega)$ is a real function in \mathbb{R} given by

$$a(\omega) = -\frac{1}{j} \int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} f(t) e^{-j\omega t} dt. \quad (65)$$

Based on this in Appendix D, we show that the PSD of the transmit signal $x(t)$ is given by

$$S_X(\omega) = \frac{\hat{P}F(\omega)}{T_{\text{avg}}} \left(1 + 2 \lim_{K \rightarrow \infty} \sum_{n=1}^{K-1} (-1)^n \left(1 - \frac{n}{K} \right) \mathbb{E}[\cos(\omega L_n)] \right), \quad (66)$$

where $n = k - j$ is the index describing the distance between two arbitrary zero-crossing instances and $L_n = T_k - T_j$ is the corresponding random variable with probability distribution

$$p_L(l_n) = \frac{\lambda^n e^{-\lambda(l_n - n\beta)} (l_n - n\beta)^{n-1}}{(n-1)!}, \quad n \geq 1, l_n \geq n\beta. \quad (67)$$

Using (67) to calculate the expectation in (66) yields

$$\mathbb{E}[\cos(\omega L_n)] = \left(\frac{\lambda}{\sqrt{\lambda^2 + \omega^2}} \right)^n \cos \left(n \left(\omega\beta + \arctan \left(\frac{\omega}{\lambda} \right) \right) \right) \leq \left(\frac{\lambda}{\sqrt{\lambda^2 + \omega^2}} \right)^n, \quad (68)$$

which can be used to upper bound the infinite sum in (66) by

$$\lim_{K \rightarrow \infty} \sum_{n=1}^{K-1} \left(1 - \frac{n}{K} \right) \left(\frac{\lambda}{\sqrt{\lambda^2 + \omega^2}} \right)^n = \frac{\lambda}{\sqrt{\lambda^2 + \omega^2} - \lambda} = c(\omega). \quad (69)$$

Hence, the PSD can be bounded as

$$S_X(\omega) \leq \frac{\hat{P}}{T_{\text{avg}}} (1 + 2c(\omega)) F(\omega). \quad (70)$$

For the sine-waveform introduced in (8), we have

$$F(\omega) = 2(1 + \cos(\omega\beta)) \left[\frac{\pi^2}{\omega(\pi^2 - \omega^2\beta^2)} \right]^2. \quad (71)$$

It can be shown that the shape of the normalized PSD $\lambda S_X(f/\lambda)$ depends only on the ratio W/λ . It is depicted in Fig. 3.

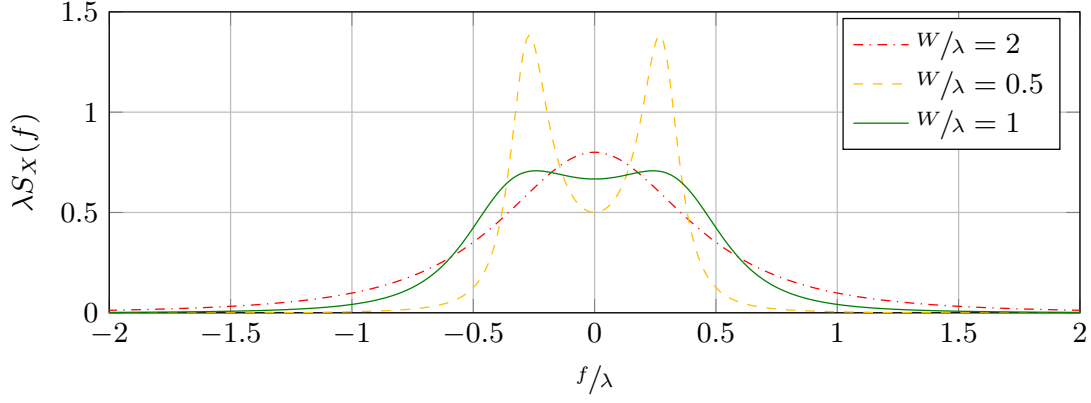


Fig. 3. Upper bound on the PSD of the transmit signal

B. Bounds on the MSE and the ACF

With the results from Section VI-A, bounds on $\sigma_{\tilde{x}}^2$ and $s''_{\tilde{x}\tilde{x}}(0)$ can be computed. For $\sigma_{\tilde{x}}^2$ we get with (61), (70) and (71)

$$\sigma_{\tilde{x}}^2 \leq \frac{1}{\pi} \int_{2\pi W}^{\infty} \frac{\hat{P}}{T_{\text{avg}}} (1 + 2c(\omega)) 2(1 + \cos(\omega\beta)) \left[\frac{\pi^2}{\omega(\pi^2 - \omega^2\beta^2)} \right]^2 d\omega. \quad (72)$$

In order to solve the integral, one further bounding step is applied. Note, that $c(\omega)$ is monotonically decreasing in ω and, hence, for all $\omega \geq 2\pi W$

$$c(\omega) \leq c(2\pi W) = \frac{\lambda}{\sqrt{\lambda^2 + 4\pi^2 W^2} - \lambda} = c_1, \quad (73)$$

such that

$$\begin{aligned} \sigma_{\tilde{x}}^2 &\leq \frac{2(1 + 2c_1)\hat{P}}{T_{\text{avg}}\pi} \int_{2\pi W}^{\infty} \left[\frac{\pi^2}{\omega(\pi^2 - \omega^2\beta^2)} \right]^2 (1 + \cos(\omega\beta)) d\omega \\ &= \frac{(1 + 2c_1)\hat{P}\beta}{2T_{\text{avg}}\pi^2} (-3\gamma - 3\log(2\pi) + 3\text{Ci}(2\pi) - \pi^2 + 4\pi\text{Si}(\pi) - \pi\text{Si}(2\pi)), \end{aligned} \quad (74)$$

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant and $\text{Si}(\cdot)$ and $\text{Ci}(\cdot)$ are the sine- and cosine-integral functions, respectively.

Furthermore, the autocorrelation function of the lowpass-distortion $\tilde{x}(t)$ is given by

$$s_{\tilde{x}\tilde{x}}(\tau) = \frac{1}{\pi} \int_{2\pi W}^{\infty} S_X(\omega) \cos(\omega\tau) d\omega, \quad (75)$$

such that for its second derivative it can be written

$$s''_{\tilde{x}\tilde{x}}(\tau) = \frac{\partial^2}{\partial \tau^2} s_{\tilde{x}\tilde{x}}(\tau) = \frac{1}{\pi} \int_{2\pi W}^{\infty} S_X(\omega) \frac{\partial^2}{\partial \tau^2} \cos(\omega\tau) d\omega, \quad (76)$$

where the exchangeability of differentiation and integration has been shown via Lebesgue's dominated convergence theorem [21, Theorem 1.34], with the dominating function

$$g(\omega) = \omega^2 S_X(\omega). \quad (77)$$

In (76) $\frac{\partial^2}{\partial \tau^2} \cos(\omega\tau)|_{\tau=0} = -\omega^2$ and since $S_X(\omega)$ is positive for all ω , an upper bound on $S_X(\omega)$ results in a lower bound on $s''_{\tilde{x}\tilde{x}}(0)$ given by

$$\begin{aligned} s''_{\tilde{x}\tilde{x}}(0) &\geq -\frac{2(1 + 2c_1)\hat{P}}{T_{\text{avg}}\pi} \int_{2\pi W}^{\infty} \frac{1 + \cos(\beta\omega)}{\left(1 - \left(\frac{\beta\omega}{\pi}\right)^2\right)^2} d\omega \\ &= -\frac{(1 + 2c_1)\hat{P}}{2T_{\text{avg}}\beta} [\pi^2 - \gamma - \log(2\pi) - \pi\text{Si}(2\pi) + \text{Ci}(2\pi)]. \end{aligned} \quad (78)$$

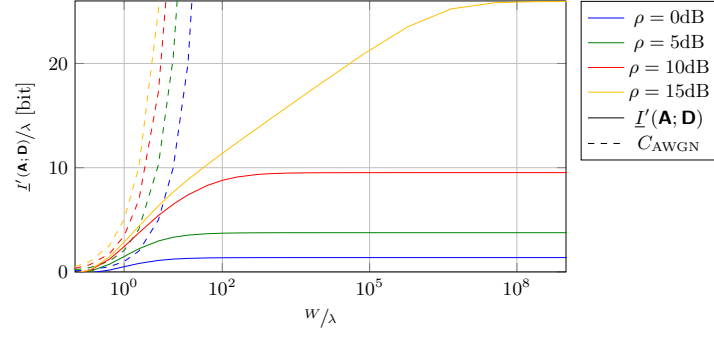


Fig. 4. Lower bound on $I'(\mathbf{A}; \mathbf{D})$ in comparison to the AWGN capacity

VII. LOWER BOUND ON THE ACHIEVABLE RATE

Substituting (2), (23), (52), (57), and (58) into (28), a lower bound on the achievable rate of the 1-bit quantized time continuous channel is given by

$$I'(\mathbf{A}; \mathbf{D}) \geq \underline{I}'(\mathbf{A}; \mathbf{D}) = \frac{2W}{2W\lambda^{-1} + 1} \left[\frac{1}{2} \log\left(\frac{e}{2\pi}\right) + \frac{1}{2} \operatorname{arcosh}\left(\frac{2\pi^2 W^2 \hat{P}}{\sigma_z^2 \lambda^2} + 1\right) + (\mu - 1) \log(\mu - 1) - \mu \log(\mu) \right] \quad (79)$$

with the equalities and inequalities (2), (33), (55), (56), (74), and (78). Fig. 4 shows the lower bound of the achievable rate given in (79) for different SNRs ρ , where both axis are normalized on λ . Hereby, the SNR is kept constant. That for constant SNR λ acts as a scaling factor for the achievable rate over W can also be seen when substituting in (79) $P/N_0 W = \rho = \text{const.}$ and $W/\lambda = k = \text{const.}$ It results

$$\underline{I}'(\mathbf{A}; \mathbf{D}) = \frac{2W}{2k + 1} \left[\frac{1}{2} \log\left(\frac{e}{2\pi}\right) + \frac{1}{2} \operatorname{arcosh}\left(2\pi^2 k^2 f_1(k, \rho) + 1\right) + f_2(k, \rho) \right] \quad (80)$$

where

$$f_1(k, \rho) = \frac{\hat{P}}{\sigma_z^2} = \frac{1 + 2k}{\frac{1}{2} + 2k} \frac{\rho}{1 + \frac{(1+2c_1(k))c_0}{2\pi^2(\frac{1}{2}+2k)} \rho} \quad (81)$$

with $c_0 = -3\gamma - 3\log(2\pi) + 3\operatorname{Ci}(2\pi) - \pi^2 + 4\pi\operatorname{Si}(\pi) - \pi\operatorname{Si}(2\pi)$ and $c_1(k) = \frac{1}{\sqrt{1+4\pi^2 k^2 - 1}}$. Furthermore,

$$f_2(k, \rho) = (\mu - 1) \log(\mu - 1) - \mu \log(\mu) \quad (82)$$

where

$$\mu = \frac{k}{\pi} \sqrt{\frac{\frac{4}{3}\pi^2 + \frac{2(1+2c_1(k))c_2}{(\frac{1}{2}+2k)} \rho}{1 + \frac{(1+2c_1(k))c_0}{2\pi^2(\frac{1}{2}+2k)} \rho}} \exp\left(-\frac{f_1(\rho, k)}{2}\right) + 1 \quad (83)$$

with $c_2 = \pi^2 - \gamma - \log(2\pi) - \pi\operatorname{Si}(2\pi) + \operatorname{Ci}(2\pi)$. Thus, $I'(\mathbf{A}; \mathbf{D})/\lambda$ is a function solely depending on W/λ .

Furthermore, the achievable rate saturates for high bandwidths W due to the limited randomness of the input signal controlled by λ . In the saturation range the average symbol duration A_k is large compared to the coherence time of the noise such that the expected number of additional zero-crossings within A_k becomes significant. Thus, the increase of the achievable rate with side information $I'(\mathbf{A}; \mathbf{D}, \mathbf{V})$ with the bandwidth W is compensated by the increase of $H'(\mathbf{V})$ representing the rate reduction due to additional zero-crossings. Moreover, in Fig. 4 it can be observed that if W is significantly smaller than λ , the lower bound becomes zero. Note that this does not mean that the achievable rate is zero, as (79) is a lower bound. However, it means that in the region of $W/\lambda \geq 0.5$, the lower bound becomes valuable. For this parameter range it has been shown in Appendix A that the assumption (A2) is valid.

For comparison also the capacity of the AWGN channel without output quantization given by

$$C_{\text{AWGN}} = W \log(1 + \rho). \quad (84)$$

This capacity is an upper bound to the capacity of the continuous-time 1-bit quantized channel studied in the present paper. It can be seen, that the lower bound is relatively tight for W/λ in the order of 1.

In order to avoid saturation of the achievable rate for the chosen input distribution, the randomness of the input signal needs to be matched to the channel bandwidth, which is achieved by allowing λ to grow linearly with W , i.e., $\lambda = \frac{1}{k}W$ with k being constant. In this case, for a given ρ , the achievable rate increases linearly with the bandwidth as depicted in Fig. 5. On a

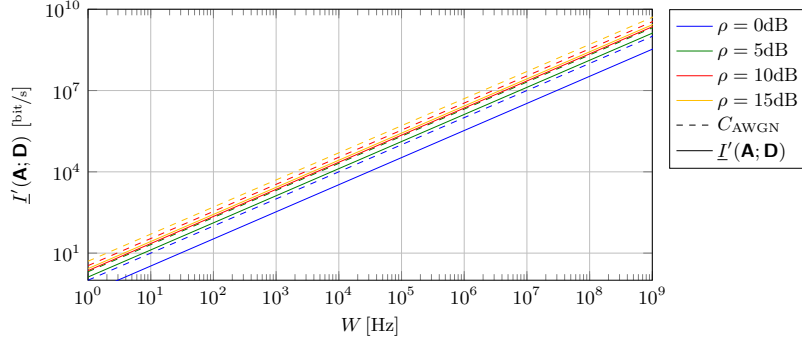


Fig. 5. Lower bound on $I'(\mathbf{A}; \mathbf{D})$ in comparison to the AWGN capacity for $W/\lambda = 1$

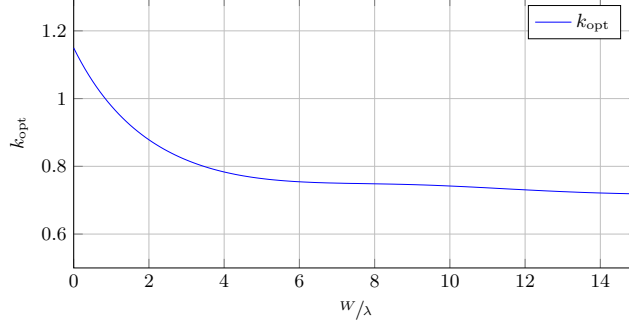


Fig. 6. Optimal ratio $k = W\lambda^{-1}$ over the SNR

logarithmic scale, two observations can be made. On the one hand, $\log C_{\text{AWGN}}$ and $\log \underline{I}'(\mathbf{A}; \mathbf{D})$ increase with the same slope. With the SNR $\rho = \text{const.}$, it can be written for the Gaussian capacity

$$\kappa_{\text{AWGN}} = \frac{\partial \log C_{\text{AWGN}}}{\partial \log W} = W \frac{1}{W \log(1 + \rho)} \log(1 + \rho) = 1. \quad (85)$$

Analogously, for the 1-bit quantized continuous time channel, it is

$$\kappa_{\underline{I}'(\mathbf{A}; \mathbf{D})} = \frac{\partial \log \underline{I}'(\mathbf{A}; \mathbf{D})}{\partial \log W} = W \frac{1}{\underline{I}'(\mathbf{A}; \mathbf{D})} \frac{\underline{I}'(\mathbf{A}; \mathbf{D})}{W} = 1, \quad (86)$$

where (86) results from (80) for $\rho = \text{const}$ and $k = W\lambda^{-1} = \text{const}$. The offset between both curves is given by

$$\Delta = \log C_{\text{AWGN}} - \log \underline{I}'(\mathbf{A}; \mathbf{D}) = \log \left(\frac{C_{\text{AWGN}}}{\underline{I}'(\mathbf{A}; \mathbf{D})} \right) \quad (87)$$

$$= \log \left[\frac{2k + 1}{2} \left(\frac{\log(1 + \rho)}{\frac{1}{2} \log\left(\frac{e}{2\pi}\right) + \frac{1}{2} \text{arcosh}(2\pi^2 k^2 f_1(k, \rho) + 1) + f_2(k, \rho)} \right) \right] \quad (88)$$

which shows, that there is a constant ratio between AWGN capacity and $\underline{I}'(\mathbf{A}; \mathbf{D})$. We evaluated the minimum of (88) w.r.t. k numerically and found that in the high SNR regime the optimal k is approximately 0.7. This is depicted in Fig. 6.

Furthermore, Fig. 7 shows the lower bound $\underline{I}'(\mathbf{A}; \mathbf{D})$ of as a function of the SNR. This shows that while increasing the SNR to infinity the lower bound $\underline{I}'(\mathbf{A}; \mathbf{D})$ saturates. Differently the AWGN capacity does not saturate. For the limiting case $\rho \rightarrow \infty$, the lower bound on the achievable rate is

$$\begin{aligned} \lim_{\rho \rightarrow \infty} \underline{I}'(\mathbf{A}; \mathbf{D}) &= \frac{1}{T_{\text{avg}}} \left[\frac{1}{2} \log\left(\frac{e}{2\pi}\right) + \frac{1}{2} \text{arcosh}\left(\frac{4\pi^4 W^2 (1 + 2W\lambda^{-1})}{\lambda^2 (1 + 2c_1)c_0} + 1\right) \right. \\ &\quad \left. - \left(2W\lambda^{-1} \sqrt{\frac{c_2}{c_0}} \exp\left(-\frac{\pi^2(1 + 2W\lambda^{-1})}{(1 + 2c_1)c_0}\right) \right) \log\left(1 + \frac{1}{2W\lambda^{-1}} \sqrt{\frac{c_0}{c_2}} \exp\left(\frac{\pi^2(1 + 2W\lambda^{-1})}{(1 + 2c_1)c_0}\right)\right) \right. \\ &\quad \left. - \log\left(2W\lambda^{-1} \sqrt{\frac{c_2}{c_0}} \exp\left(-\frac{\pi^2(1 + 2W\lambda^{-1})}{(1 + 2c_1)c_0}\right) + 1\right) \right]. \end{aligned} \quad (89)$$

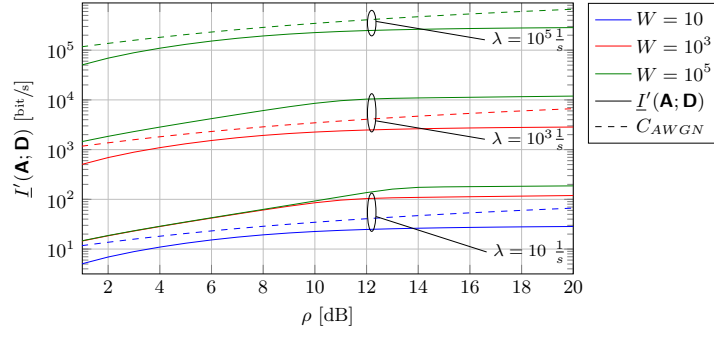


Fig. 7. Lower bound on $I'(\mathbf{A}; \mathbf{D})$ depending on the SNR ρ

APPENDIX A NUMBER OF ZERO-CROSSINGS WITHIN A TRANSITION INTERVAL

We want to verify the assumption (A2) that within the interval $[T_k - \beta/2, T_k + \beta/2]$ with very high probability only one zero-crossing occurs. This is a curve crossing problem depending on the stationary Gaussian process $\hat{n}(t)$ and the deterministic waveform $f(t)$ within the given interval. An equivalent way of looking at this problem is to study the zero-crossings of a non-stationary Gaussian process $q(t) = \hat{n}(t) - \psi(t)$, where $\psi(t)$ is the deterministic curve to be crossed by the random process. For this purpose we define the transition interval $\mathbb{Y} = [0, \beta]$, where $y \in \mathbb{Y}$ is the time variable within the transition interval. Then the deterministic function

$$\psi(y) = -f\left(y - \frac{\beta}{2}\right) \quad (90)$$

depends on the waveform $f(t)$ of the transition. For the sine transition in (7) it is given by

$$\psi(y) = \cos\left(\frac{\pi}{\beta}y\right). \quad (91)$$

The process $q(t)$ has a zero-crossing, if and only if $\hat{n}(y) = \psi(y)$. For the number of crossings $N_T(\psi)$ of a curve ψ by a stationary Gaussian processes in the time interval T holds [22]

$$\mathbb{E}[N_T(\psi)] = \sqrt{-s''(0)} \int_0^T \varphi(\psi(y)) \left[2\varphi\left(\frac{\psi'(y)}{\sqrt{-s''(0)}}\right) + \frac{\psi'(y)}{\sqrt{-s''(0)}} \left(2\Phi\left(\frac{\psi'(y)}{\sqrt{-s''(0)}}\right) - 1 \right) \right] dy, \quad (92)$$

where $s(\tau)$ is the ACF of the Gaussian Process, ' denotes the derivative in time, i.e., w.r.t. y , and φ and Φ are the Gaussian density and distribution function of $q(t)$, respectively. The variance of the number of zero-crossings is given by [22]

$$\begin{aligned} \text{Var}(N_T(\psi)) &= \mathbb{E}[N_T(\psi)] - \mathbb{E}^2[N_T(\psi)] \\ &+ \int_0^T \int_0^T \int_{\mathbb{R}} |\mathbf{q}'_{t_1} - \psi'_{t_1}| |\mathbf{q}'_{t_2} - \psi'_{t_2}| \phi_{t_1, t_2}(\psi_{t_1}, \mathbf{q}'_{t_1}, \psi_{t_2}, \mathbf{q}'_{t_2}) d\mathbf{q}'_{t_1} d\mathbf{q}'_{t_2} dt_1 dt_2, \end{aligned} \quad (93)$$

where ϕ is the multivariate normal distribution of $\mathbf{q}(t_1)$, $\mathbf{q}'(t_1)$, $\mathbf{q}(t_2)$, and $\mathbf{q}'(t_2)$ with covariance matrix Σ

$$\Sigma = \begin{Bmatrix} s(0) & 0 & s(\tau) & s'(\tau) \\ 0 & -s''(0) & -s'(\tau) & -s''(\tau) \\ s(\tau) & -s'(\tau) & s(0) & 0 \\ s'(\tau) & -s''(\tau) & 0 & -s''(0) \end{Bmatrix}. \quad (94)$$

The equations (92) and (93) are evaluated and depicted in Fig. 8. For W/λ in the order of 1, the expectation of the number of zero-crossings converges to 1 for $\text{SNR} \geq 5$ dB while at the same time the variance converges to 0. Hence, for $\text{SNR} \geq 5$ dB almost surely exactly one zero-crossing exists in every transition interval. For $W/\lambda \ll 1$, the lower bound in (79) becomes zero and, hence, the validity of the assumption is not relevant.

APPENDIX B VALIDITY OF THE GAUSSIAN APPROXIMATION

In order to quantify the high-SNR region for which the approximation of $p_S(s)$ in (34) by the Gaussian density in (35) is valid, the variances of both densities have been evaluated and compared numerically. The results are depicted in Fig. 9 and show a convergence of the variances in the relevant area $\frac{W}{\lambda} \geq 1$ for SNRs larger 6 dB. Comparing the variances is sufficient for our purpose as the further bounding of $I'(\mathbf{A}; \mathbf{D}, \mathbf{V})$ is solely based on the variance of a Gaussian random process with equal covariance matrix.

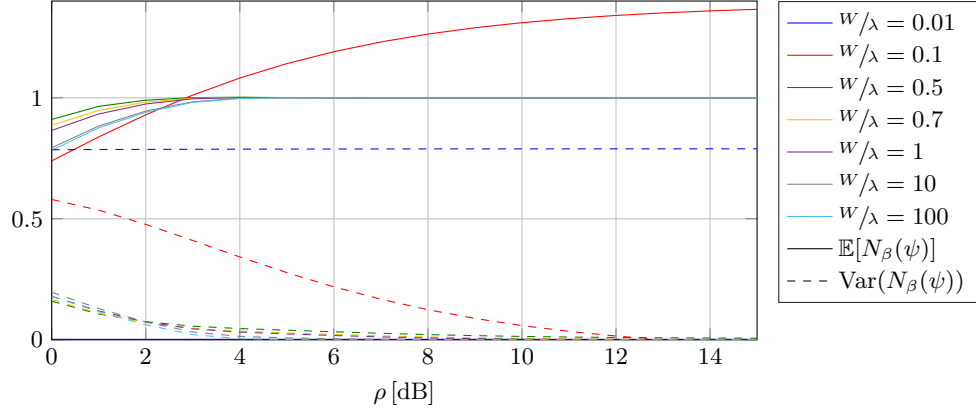


Fig. 8. Expectation and variance of the number of zero-crossings in the transition interval $[T_k - \beta/2, T_k + \beta/2]$, $\psi(y) = \cos\left(\frac{\pi}{\beta}y\right)$

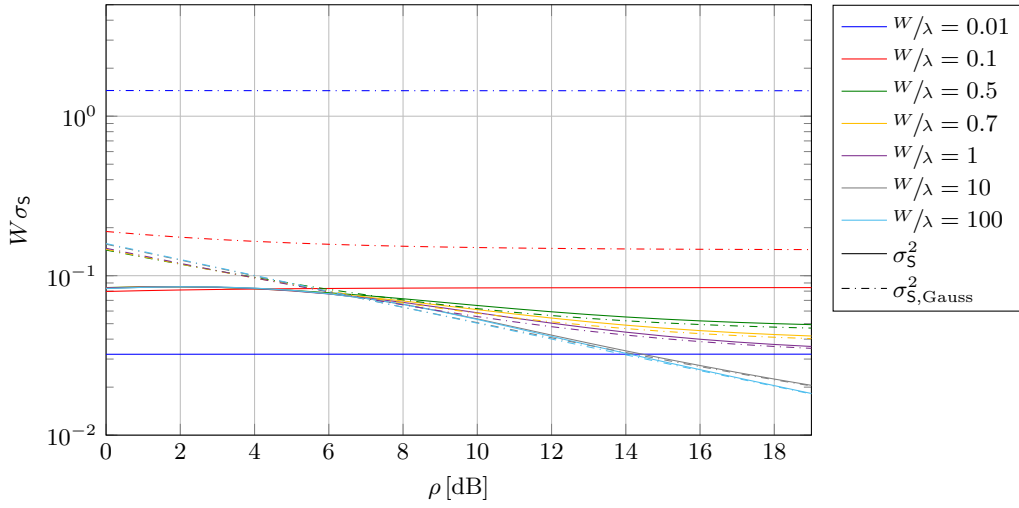


Fig. 9. Variances of the original distribution σ_S^2 (34) and the Gaussian approximation $\sigma_{S,\text{Gauss}}^2$ (35)

APPENDIX C THE ENTROPY OF V_k

The entropy maximizing distribution for a discrete, positive random variable with given mean μ is the geometric distribution [20]

$$p_i = Cq^i, \quad i \in \mathbb{N} \quad (95)$$

where $C = \frac{1}{\mu-1}$ and $q = \left(\frac{\mu-1}{\mu}\right)$. Thus, the entropy of V_k becomes

$$\begin{aligned} H(V) &= -\sum_{i=1}^{\infty} p_i \log p_i = -\sum_{i=1}^{\infty} Cq^i \log Cq^i \\ &= -C \log(C) \sum_{i=1}^{\infty} q^i - C \log(q) \sum_{i=1}^{\infty} q^i i \\ &= -C \log(C) \frac{q}{1-q} - C \log(q) \frac{q}{(q-1)^2} \\ &= -\frac{1}{\mu-1} \log\left(\frac{1}{\mu-1}\right) \frac{\frac{\mu-1}{\mu}}{1-\frac{\mu-1}{\mu}} - \frac{1}{\mu-1} \log\left(\frac{\mu-1}{\mu}\right) \frac{\frac{\mu-1}{\mu}}{\left(\frac{\mu-1}{\mu}-1\right)^2} \\ &= (1-\mu) \log(\mu-1) + \mu \log \mu. \end{aligned} \quad (96)$$

APPENDIX D
POWER SPECTRAL DENSITY OF THE TRANSMIT SIGNAL

When separating the spectrum $X(\omega)$ of $x(t)$ in (63) into real and imaginary part, we obtain

$$X(\omega) = 2\pi\sqrt{\hat{P}}\delta(\omega) + \sqrt{\hat{P}}\sum_{k=1}^K(-1)^{k+1}\gamma_k + j\sqrt{\hat{P}}\sum_{k=1}^K(-1)^{k+1}\nu_k \quad (97)$$

where we introduced the notation

$$\begin{aligned} \gamma_k &= -\Re\{G(\omega)e^{-j\omega T_k}\} \\ &= \frac{1}{\omega}[(1 + \cos(\omega\beta))\sin(\omega T_k) + \sin(\omega\beta)\cos(\omega T_k)] + a(\omega)\left[\cos\left(\frac{\omega\beta}{2}\right)\sin(\omega T_k) + \sin\left(\frac{\omega\beta}{2}\right)\cos(\omega T_k)\right] \end{aligned} \quad (98)$$

$$\begin{aligned} \nu_k &= -\Im\{G(\omega)e^{-j\omega T_k}\} \\ &= \frac{1}{\omega}[(1 + \cos(\omega\beta))\cos(\omega T_k) - \sin(\omega\beta)\sin(\omega T_k)] + a(\omega)\left[\cos\left(\frac{\omega\beta}{2}\right)\cos(\omega T_k) - \sin\left(\frac{\omega\beta}{2}\right)\sin(\omega T_k)\right]. \end{aligned} \quad (99)$$

We thus can write for the expectation in (62)

$$\mathbb{E}[|X(\omega)|^2] = 4\pi^2\hat{P}\delta^2(\omega) + 2\pi\hat{P}\delta(\omega)\sum_{k=1}^K(-1)^{k+1}\mathbb{E}[\gamma_k] + \hat{P}\sum_{k=1}^K\sum_{j=1}^K(-1)^{k+j}\mathbb{E}[\gamma_k\gamma_j + \nu_k\nu_j]. \quad (100)$$

In the second term of (100) the Dirac-function can be used to relieve the computation of the expectation, yielding

$$\mathbb{E}[\delta(\omega)\gamma_k] = \mathbb{E}\left[\lim_{\omega \rightarrow 0} \gamma_k\right] = 2T_{\text{avg}} + \beta. \quad (101)$$

For the third term of (100) it holds, that due to the duality of γ_k and ν_k , the remaining function is an even function and only depending on the difference of T_k and T_j

$$\mathbb{E}[\gamma_k\gamma_j + \nu_k\nu_j] = \mathbb{E}[\cos(\omega(T_k - T_j))]F(\omega) \quad (102)$$

where

$$F(\omega) = \frac{2(1 + \cos(\omega\beta))}{\omega^2} + a^2(\omega) + \frac{2a(\omega)}{\omega}\left((1 + \cos(\omega\beta))\cos\left(\frac{\omega\beta}{2}\right) + \sin(\omega\beta)\sin\left(\frac{\omega\beta}{2}\right)\right). \quad (103)$$

Due to (101) being independent of k and exploiting the fact that the cosine is an even function, we get for the PSD of the transmit signal in (62)

$$S_X(\omega) = \lim_{K \rightarrow \infty} \frac{\hat{P}}{KT_{\text{avg}}} \left(4\pi^2\delta^2(\omega) + 2\pi(2T_{\text{avg}} + \beta)\sum_{k=1}^K(-1)^{k+1} + F(\omega)\sum_{k=1}^K\sum_{j=1}^K(-1)^{k+j}\mathbb{E}[\cos(\omega(T_k - T_j))] \right) \quad (104)$$

$$= \frac{\hat{P}F(\omega)}{T_{\text{avg}}} \lim_{K \rightarrow \infty} \frac{1}{K} \left(\sum_{k=1}^K(-1)^{2k} + \sum_{k=1}^K\sum_{\substack{j=1 \\ j \neq k}}^K(-1)^{k+j}\mathbb{E}[\cos(\omega(T_k - T_j))] \right) \quad (105)$$

$$= \frac{\hat{P}F(\omega)}{T_{\text{avg}}} \left(1 + \lim_{K \rightarrow \infty} 2\sum_{n=1}^{K-1}(-1)^n\left(1 - \frac{n}{K}\right)\mathbb{E}[\cos(\omega L_n)] \right) \quad (106)$$

where $n = k - j$ is the index describing the distance between two arbitrary zero-crossing instances and $L_n = T_k - T_j$ is the corresponding random variable with probability distribution

$$p_L(l_n) = \frac{\lambda^n e^{-\lambda(l_n - n\beta)}(l_n - n\beta)^{n-1}}{(n-1)!} \quad n \geq 1, l_n \geq n\beta. \quad (107)$$

This results from the fact that L_n is the sum of n consecutive input symbols

$$L_n = \sum_{i=1}^n A_{k+i}. \quad (108)$$

As the input is i.i.d., it holds

$$p(A_k + 1, \dots, A_{k+n}) = p(A_1, \dots, A_n) = \prod_{i=1}^n p(A_i). \quad (109)$$

From (108) and (109) we obtain (107).

REFERENCES

- [1] L. Landau and G. Fettweis, "Information rates employing 1-bit quantization and oversampling at the receiver," in *Proc. of the IEEE Int. Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, Toronto, June 2014, pp. 219–223.
- [2] —, "On reconstructable ASK-sequences for receivers employing 1-bit quantization and oversampling," in *Proc. of the IEEE Int. Conference on Ultra-WideBand (ICUWB)*, Paris, Sept 2014, pp. 180–184.
- [3] L. Landau, M. Dörpinghaus, and G. Fettweis, "Communications employing 1-bit quantization and oversampling at the receiver: Faster-than-Nyquist signaling and sequence design," in *Proc. of the IEEE Int. Conf. on Ubiquitous Wireless Broadband (ICUWB)*, Montreal, Oct 2015.
- [4] S. Bender, L. Landau, M. Dörpinghaus, and G. Fettweis, "Communication with 1-bit quantization and oversampling at the receiver: Spectral constrained waveform optimization," in *Proc. of the IEEE Int. Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, Edinburgh, July 2016.
- [5] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 623–656, 1948.
- [6] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 4–18, 1996.
- [7] E. N. Gilbert, "Increased information rate by oversampling," *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1973–1976, 1993.
- [8] S. Shamai, "Information rates by oversampling the sign of a bandlimited process," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1230–1236, 1994.
- [9] W. Zhang, "A general framework for transmission with transceiver distortion and some applications," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 384–399, 2012.
- [10] R. G. Gallager, "Sequential decoding for binary channels with noise and synchronization errors," Massachusetts Institute of Technology: Lincoln Laboratory, Tech. Rep., 1961.
- [11] K. S. Zigangirov, "Sequential decoding for a binary channel with drop-outs and insertions," *Problems Inform. Transmission*, vol. 5, pp. 17–22, 1969.
- [12] D. Fertonani, T. Duman, and M. Erden, "Bounds on the capacity of channels with insertions, deletions and substitutions," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 2–6, January 2011.
- [13] S. Diggavi, M. Mitzenmacher, and H. Pfister, "Capacity upper bounds for deletion channels," in *Proc. of the Int. Symposium on Information Theory*, 2007, pp. 1716–1720.
- [14] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd edition. New York: Wiley & Sons, 2006.
- [15] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications*. Berkeley, CA, U.S.A.: Univ. Calif. Press, 1958.
- [16] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.
- [17] M. Kac, "On the average number of real roots of a random algebraic equation," *Bull. Amer. Math. Soc.*, vol. 49, no. 4, pp. 314–320, 04 1943.
- [18] S. O. Rice, "Mathematical analysis of random noise," *Bell System Technical Journal*, vol. 23, no. 3, pp. 282–332, 1944.
- [19] H. Cramer and M. R. Leadbetter, *Stationary and Related Stochastic Processes*, 1967.
- [20] J. Kapur, *Maximum-entropy Models in Science and Engineering*. Wiley Eastern, New Dehli, 1989.
- [21] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill Book Co., 1987.
- [22] M. F. Kratz, "Level crossings and other level functionals of stationary Gaussian processes," *Probab. Surveys*, vol. 3, pp. 230–288, 2006.